

KLASIFIKASI PRESTASI AKADEMIK
MAHASISWA FKI UMS
MENGGUNAKAN METODE *DECISION TREE*

Makalah

Program Studi Teknik Informatika

Fakultas Komunikasi Dan Informatika



Diajukan Oleh :

Setyawan

Yusuf Sulistyo Nugroho, ST., M. Eng.

PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA

JANUARI 2014

HALAMAN PENGESAHAN

Publikasi Ilmiah dengan Judul

**KLASIFIKASI PRESTASI AKADEMIK
MAHASISWA FKI UMS
MENGUNAKAN METODE *DECISION TREE***

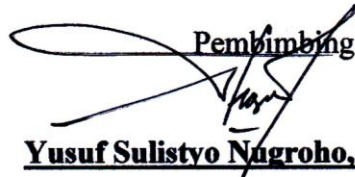
Yang dipersiapkan dan disusun oleh :

Setyawan

L200100155

Telah Disetujui Pada:

Hari : Sabtu
Tanggal : 25 Januari 2014


Pembimbing

Yusuf Sulistyono, S.T., M.Eng

Nik : 100.1197

Publikasi Ilmiah Ini Telah Diterima Sebagai Salah Satu Persyaratan

Untuk Memperoleh Gelar Sarjana

Tanggal.....

Mengetahui,

Ketua Program Studi

Teknik Informatika

Dr. Heru Supriyono, M.Sc.

Nik : 970

KLASIFIKASI PRESTASI AKADEMIK
MAHASISWA FKI UMS
MENGGUNAKAN METODE *DECISION TREE*

Setyawan, Yusuf Sulistyo Nugroho
Teknik Informatika, Fakultas Komunikasi dan Informatika,
Universitas Muhammadiyah Surakarta
Email : wanzetya@gmail.com

Abstraction

Since the establishment of the FKI UMS until now, FKI UMS has students as much as 2358 people. So it has very much academic data. The data needs processed in order to make a useful information using decision tree method, and the attributes selection is based on information gain to provide a strategic plan for FKI UMS in looking for new students and increase student academic achievement. The steps taken are collecting data from the Bureau of Academic Administration (BAA), FKI UMS , and Information Technology Dept. which all are in the scope of Muhammadiyah University of Surakarta . Then, analyzing the data using Rapid Miner and application manual calculation method using the Decision Tree algorithm to find the value of C 5.0 entropy and information gain . The result of classification shows some attributes have a high dominance in influencing GPA such as IPA concentration. If the gender is male, as assistant, from any department at the high school, then he is predicted to have a satisfactory GPA . However, if the gender is male, and the region of the school is from Surakarta residency, the predicted of his GPA is less satisfactory, if not an assistant.

Keywords : Academic Achievement, Classification, Data Mining , Decision Tree, Information Gain.

Abstraksi

Sejak berdirinya UMS FKI sampai sekarang, FKI UMS memiliki mahasiswa sebanyak 2358 orang sehingga memiliki data akademik yang sangat banyak. Data tersebut perlu diolah agar bermanfaat dengan cara klasifikasi data mining metode *decision tree*, serta pemilihan atributnya berdasarkan *information gain* dengan tujuan memberikan rencana strategis bagi FKI UMS dalam mencari mahasiswa baru serta meningkatkan prestasi akademik mahasiswa. Adapun langkah yang dilakukan yaitu pengumpulan data dari Biro Administrasi Akademik (BAA), FKI UMS, serta *Information Technology* yang semua berada di lingkup Universitas Muhammadiyah Surakarta. Kemudian menganalisa data

menggunakan aplikasi *Rapid Miner* serta perhitungan manual dengan metode *Decision Tree* menggunakan algoritma C 5.0 dengan mencari nilai *entropy* serta *information gain*. Hasil klasifikasi menunjukkan beberapa atribut mempunyai dominasi yang tinggi terhadap nilai IPK seperti jurusan di SMA adalah IPA. Apabila jenis kelamin pria, menjadi asisten, dari jurusan apapun di SMA nya, maka diprediksi memiliki IPK yang memuaskan. Namun apabila berjenis kelamin pria, dari daerah serta asal sekolah dari karesidenan Surakarta, maka diprediksi IPK nya kurang memuaskan, apabila tidak menjadi asisten.

Kata kunci : *Data Mining, Decision Tree, Information Gain*, Klasifikasi, Prestasi Akademik

I. PENDAHULUAN

Sejak berdirinya UMS FKI sampai sekarang, FKI UMS memiliki mahasiswa sebanyak 2358 orang baik yang lulus maupun belum lulus. Padahal FKI sendiri berdiri mulai pada tahun 2006. Sampai sekarang FKI sudah meluluskan mahasiswa sekitar 388 mahasiswa. Bagi fakultas yang baru berumur 7 tahun dengan jenjang perkuliahan S1 minimal 3.5 tahun, jumlah tersebut merupakan suatu keberhasilan yang membanggakan.

Dengan membludaknya mahasiswa, maka data – data akademik yang dimiliki juga semakin bertambah banyak. Apabila tidak diolah dengan baik dan benar, hanya akan menjadi tumpukan data yang tidak berguna dan tidak bermanfaat.

Padahal data tersebut bisa dijadikan sebagai sumber data strategis yang sangat bermanfaat bagi Fakultas Komunikasi dan Informatika UMS.

Berdasarkan permasalahan tersebut, maka dalam penelitian ini penulis akan mengklasifikasikan prestasi akademik yang diperoleh mahasiswa menggunakan *data mining* metode *decision tree*, serta pemilihan atributnya berdasarkan *information gain* menggunakan *algoritma C. 5.0*.

Tujuan dari penelitian ini adalah untuk mengklasifikasi prestasi akademik mahasiswa yang diperoleh para mahasiswa FKI UMS menggunakan *data mining* serta memberikan rencana strategis bagi Fakultas Komunikasi dan Informatika berdasarkan interpretasi hasil penelitian, dalam hal prioritas daerah

asal atau asal sekolah untuk calon mahasiswa baru, pembimbingan dan pengarahan yang lebih intensif bagi mahasiswa IPKnya kurang serta skala prioritas untuk memilih calon mahasiswa FKI UMS berdasarkan jurusan di SMA agar menghasilkan indeks prestasi mahasiswa yang memuaskan.

II. TINJAUAN PUSTAKA

1. Klasifikasi

Klasifikasi adalah proses pencarian sekumpulan model yang menggambarkan dan membedakan kelas data dengan tujuan agar model tersebut dapat digunakan untuk memprediksi kelas dari suatu obyek yang belum diketahui kelasnya. (Herman, et al., 2012)

Klasifikasi data menurut jenis data terdiri dari :

a. Data hitung .

Data hitung adalah hasil dari penghitungan atau jumlah tertentu. Misalnya jumlah mahasiswa dalam suatu kelas.

b. Data ukur.

Data ukur adalah data yang menunjukkan ukuran mengenai nilai sesuatu. Misalnya angka

atau huruf tertentu yang diberikan dosen kepada mahasiswa.

Klasifikasi data menurut sumber data adalah :

a. Data internal.

Data internal adalah data yang asli, artinya data sebagai hasil observasi yang dilakukan sendiri.

b. Data eksternal.

Data eksternal adalah data hasil observasi orang lain.

2. Indeks Prestasi Kumulatif (IPK)

IPK adalah angka yang menunjukkan prestasi atau keberhasilan studi mahasiswa dari semester pertama sampai dengan semester terakhir yang telah ditempuh secara kumulatif. IPK digunakan untuk :

a. Menentukan beban studi yang

dapat diambil mahasiswa pada semester berikutnya.

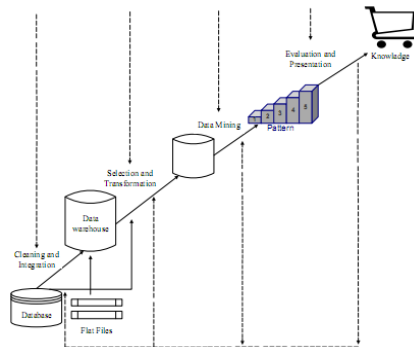
b. Evaluasi Akademik per semester.

c. Evaluasi hasil studi pada akhir program.

3. Data Mining

Data mining adalah suatu proses untuk menemukan suatu

pengetahuan atau informasi yang berguna dari data berskala besar. Sering juga disebut sebagai bagian proses KDD (*Knowledge Discovery in Databases*) (Santosa, 2007).



Gambar 1. Tahap Tahap Data Mining

a. Pembersihan data.

Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten.

b. Integrasi data.

Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru.

c. Seleksi Data.

Data yang sesuai untuk dianalisis yang akan diambil dari database.

d. Transformasi data.

Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam *data mining*.

e. Proses mining

Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan tersembunyi dari data.

f. Evaluasi pola.

Untuk mengidentifikasi pola - pola yang menarik kedalam *knowledge based* yang ditemukan.

g. Presentasi pengetahuan.

Merupakan visualisasi dan penyajian mengenai metode yang digunakan untuk memperoleh pengetahuan.

4. *Decision Tree*

Decision Tree adalah sebuah struktur pohon, dimana setiap *node* pohon merepresentasikan atribut yang telah diuji, setiap cabang merupakan suatu pembagian hasil uji, dan *node* daun merepresentasikan kelompok kelas tertentu. Level *node* teratas dari sebuah *decision Tree* adalah *node* akar (*root*) yang biasanya berupa atribut yang paling memiliki pengaruh terbesar pada suatu kelas tertentu. Pada umumnya *decision Tree* melakukan strategi pencarian secara *top-down* untuk solusinya. Pada proses mengklasifikasi data

yang tidak diketahui, nilai atribut akan diuji dengan cara melacak jalur dari *node* akar (*root*) sampai *node* akhir (daun) dan kemudian akan diprediksi kelas yang dimiliki oleh suatu data baru tertentu.(Hastuti, 2012)

5. *Rapid Miner*

Rapid miner merupakan *software tool Open Source* untuk *data Mining*. *Rapid miner* dioperasikan pada sebuah lingkungan untuk *machine learning, data mining, text mining* dan *predictive analytics*.

Desain proses analisa pada *Rapid Miner* sudah memenuhi :

a. Fleksibilitas.

Fleksibilitas karena sangat fleksibel untuk mendefinisikan proses analisa secara visual dengan GUI.

b. Fungsionalitas

Fungsionalitas karena meliputi lebih dari 500 fungsionalitas *data mining* dalam bentuk operator-operator.

c. Skalabilitas

1) Mulai versi 4.6 fokus utama pada skalabilitas untuk data ukuran besar.

2) Konsep *view* untuk data mirip seperti database.

3) Transformasi data *on-the-fly* tanpa *copy* .

4) 100 juta data set bukanlah data yang besar.

d. Format data

1) Terhubung sangat baik dengan berbagai macam sumber data: *Oracle, IBM DB2, Microsoft SQL Server,MySQL, PostgreSQL, Ingres, Excel, Access, SPSS,CSV files* dan berbagai format lain.

2) Bersama-sama dengan operator-operator untuk data *preprocessing*, bisa sebagai *tool ETL(extraction, transformation, loading)* dengan hasil yang menakjubkan.(Romi, 2012)

III. METODOLOGI PENELITIAN

1. Pengumpulan Data

Analisa pengumpulan data yang akan digunakan dalam pembangunan *data mining*. Menentukan data-data yang akan diperlukan dalam *data mining* setelah menganalisa identifikasi masalah. Dalam

mengklasifikasi indeks prestasi akademik mahasiswa maka diperlukan data-data semua mahasiswa baik yang sudah lulus maupun yang belum lulus yaitu:

Tabel 1. Daftar Atribut

Atribut	Isi dalam Atribut-Atribut
Jurusan asal sekolah	IPA, IPS dan Lain
Gender	Pria dan Wanita
Daerah asal mahasiswa	Surakarta dan Luar Surakarta
Asal sekolah	Surakarta dan Luar Surakarta
Rata – rata SKS (satuan kredit semester)	$SKS \leq 18$ dan $SKS > 18$
Rata – rata MK (Mata kuliah)	$MK \leq 7$ dan $MK > 7$
Asisten Lab	Ya dan Tidak
Indeks Prestasi Akademik	Memuaskan dan Kurang

2. Penentuan sampel

Dalam penelitian ini jumlah sampel yang harus diambil dari jumlah mahasiswa FKI secara keseluruhan digunakan rumus slovin dengan nilai toleransi maksimal $e = 5\%$. (Umar, 2004)

$$n = \frac{N}{1 + N e^2}$$

.. (1)

Bahwa :

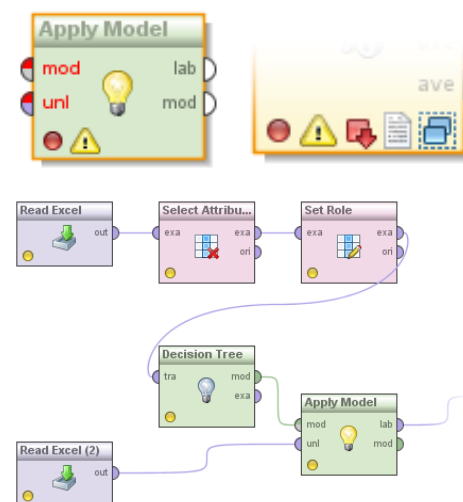
n = jumlah sampel

N = jumlah populasi

e = nilai toleransi

3. Penggunaan Rapid Miner 5.

Pada penerapan *data mining* ini dibantu menggunakan aplikasi Rapid Miner 5. Rapid Miner 5 merupakan aplikasi data mining *open source* yang terdiri atas koleksi algoritma *machine learning* yang dapat digunakan untuk melakukan generalisasi dari sekumpulan data percobaan. Sebuah operator bisa disambungkan melalui *port* masukan (kiri) dan *port* keluaran (kanan).



Gambar 2. Indikator status dari operator

- a. Lampu status: merah (tak tersambung), kuning (lengkap tetapi belum dijalankan), hijau (sudah berhasil dijalankan).
- b. Segitiga *warning*: bila ada pesan status.
- c. *Breakpoint*: bila ada *breakpoint* sebelum/sesudahnya.
- d. *Comment*: bila ada komentar.
- e. *Subprocess*: bila mempunyai *subprocess*.

Sebuah proses analisa yang terdiri dari beberapa operator. Warna aliran data menunjukkan tipe obyek yang dilewatkan. (Bowo, 2011)

4. Penggunaan Metode Decision Tree

Decision tree adalah salah satu metode untuk pengklasifikasian data. Dalam metode *decision tree*, model dipresentasikan dalam bentuk *tree*. Pohon (*Tree*) adalah sebuah struktur data yang terdiri dari simpul (*node*) dan rusuk (*edge*). Simpul pada sebuah pohon terdiri dari 3:

- a. Simpul akar (*root node*).
- b. Simpul percabangan/internal (*branch/internal node*).
- c. Simpul daun (*leaf node*).

Hal yang harus dilakukan dalam metode *decision tree* adalah

menghitung *entropy* dan *information gain*. (Ranny, Budi, 2012)

Entropy adalah ukuran kemurnian, semakin murni, semakin homogen, semakin rendah nilainya.

$$Entropi(y) = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots - p_n \log_2 p_n \quad (2)$$

Information Gain adalah pengurangan *entropy* disebabkan oleh partisi berdasarkan suatu atribut.

$$gain(y, A) = entropi(y) - \sum_{c \in \text{Nilai}(A)} \frac{y_c}{y} entropi(y_c) \quad (3)$$

IV. HASIL DAN PEMBAHASAN

1. Pengambilan jumlah sampel

Untuk mendapatkan jumlah sampel dalam penelitian ini menggunakan rumus *slovin*. Dari jumlah seluruh mahasiswa FKI UMS sejumlah 2358 mahasiswa maka jumlah sampel yang diambil yaitu:

$$n = 2358 / 1 + 2358 \times (0,05)^2$$

$$n = 2358 / 1 + 2358 \times 0,0025$$

$$n = 2358 / 1 + 5,895$$

$$n = 2358 / 6,895$$

$$n = 341,9869 \text{ mahasiswa}$$

Jadi dibulatkan menjadi 342 mahasiswa yang digunakan bahan sampel untuk data uji.

2. Implementasi Dengan Decision Tree

Dari data lulusan dapat dihitung dengan metode *decision tree*. Dari rumus pencarian *entropy* dan *information gain* penulis akan mencoba perhitungan manual untuk mengklasifikasi data prestasi akademik mahasiswa FKI UMS untuk 1 *node* teratas menggunakan *data training*. Penghitungan *data training* diatas berdasarkan data test:

Asumsi :

Y = Indeks Prestasi Akademik

X₁ = Jurusan SMA/SMK

X₂ = Gender

X₃ = Daerah / Alamat

X₄ = Asal sekolah IPK

X₅ = Rata-rata SKS

X₆ = Rata-rata mata kuliah

X₇ = Asisten Lab. FKI

a. Perhitungan untuk mencari *root node*/simpul akar

- 1) Jika perbandingan dua kelas, rasionya sama maka nilai entropinya 1.
- 2) Jika satu set hanya terdiri dari satu kelas maka entropinya 0.
- 3) Berdasarkan data pada *data training*, pada variabel y (yaitu variabel *dependent*) dalam penelitian ini variabel *dependent* adalah atribut IPK.

a) 153 KURANG.

b) 188 MEMUASKAN, sehingga diperoleh:

$$Entropi(153,188) = -\frac{153}{341} \log_2 \left(\frac{153}{341} \right) - \frac{188}{341} \log_2 \left(\frac{188}{341} \right) = 0,992$$

4) Berdasarkan data pada *data training*, pada variabel X1 (yaitu variabel *independent*) dalam penelitian ini variabel X1 adalah atribut Jurusan.

a) IPA ada 62 KURANG dan 113 MEMUASKAN.

b) IPS ada 82 KURANG dan 67 MEMUASKAN.

c) LAIN ada 9 KURANG dan 8 MEMUASKAN, sehingga diperoleh:

$$Entropi IPA(62,113) = -\frac{62}{175} \log_2 \left(\frac{62}{175} \right) - \frac{113}{175} \log_2 \left(\frac{113}{175} \right) = 0,938$$

$$Entropi IPS(82,67) = -\frac{82}{149} \log_2 \left(\frac{82}{149} \right) - \frac{67}{149} \log_2 \left(\frac{67}{149} \right) = 0,993$$

$$Entropi LAIN(9,8) = -\frac{9}{17} \log_2 \left(\frac{9}{17} \right) - \frac{8}{17} \log_2 \left(\frac{8}{17} \right) = 0,998$$

$$Entropi TOTAL(62,113),(82,67),(9,8) = \frac{175}{341}(0,938) + \frac{149}{341}(0,993) + \frac{17}{341}(0,998) = 0,965$$

$$Gain jurusan = 0,992 - \frac{175}{341}(0,938) - \frac{149}{341}(0,993) - \frac{17}{341}(0,998) = 0,028$$

Setelah dilakukan perhitungan secara menyeluruh dari atribut X1 sampai X7 didapatkan nilai masing – masing *information gain* yaitu :

$$Gain jurusan = 0,992 - \frac{175}{341}(0,938) - \frac{149}{341}(0,993) - \frac{17}{341}(0,998) = 0,028$$

$$Gain gender = 0,992 - \frac{211}{341}(0,993) - \frac{130}{341}(0,862) = 0,050$$

$$\begin{aligned}
\text{Gain daerah} &= 0,992 - \frac{92}{341}(0,999) - \frac{249}{341}(0,989) = 0,001 \\
\text{Gain Asal Sekolah} &= 0,992 - \frac{84}{341}(0,980) - \frac{257}{341}(0,995) = 0,001 \\
\text{Gain SKS} &= 0,992 - \frac{170}{341}(1,000) - \frac{171}{341}(0,970) = 0,008 \\
\text{Gain Rata - rata MatKul} &= 0,992 - \frac{141}{341}(0,992) - \frac{200}{341}(0,993) = 0,000 \\
\text{Gain Asisten} &= 0,992 - \frac{39}{341}(0,679) - \frac{302}{341}(0,999) = 0,030
\end{aligned}$$

Dari Perhitungan yang telah dilakukan maka yang menempati *root node* adalah **Gender** karena memiliki **information gain tertinggi**. Langkah selanjutnya adalah mencari nilai *information gain* dari seluruh atribut yang digunakan dalam penelitian ini.

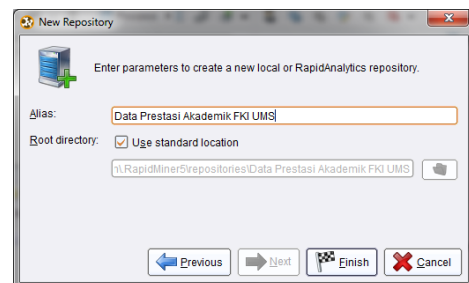
3. Implementasi dalam Rapid

Miner 5

Rapid miner 5 merupakan *software tool Open Source* untuk *Data Mining*. *Rapid miner 5* menyediakan prosedur *data mining* dan *machine learning* termasuk: ETL (*extraction, transformation, loading*), *data preprocessing*, visualisasi, *modelling* dan evaluasi. Data akademik mahasiswa Fakultas Komunikasi dan Informatika UMS baik sudah lulus maupun belum lulus yang diperoleh dari Biro Administrasi Akademik UMS, Information Technology UMS serta dari Fakultas FKI yang sudah dimiliki kemudian diolah melalui aplikasi *Rapid Miner 5* untuk mengetahui klasifikasi prestasi

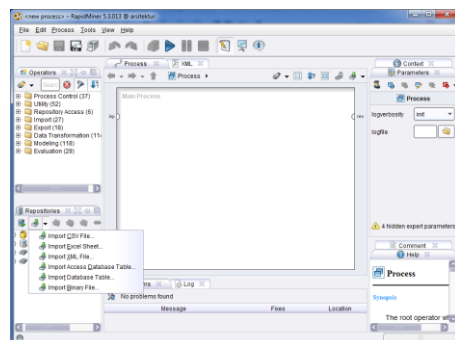
akademik mahasiswa FKI UMS menggunakan *decision tree*.

Menjalankan *Rapid Miner 5* pertama kali, membuat repositori baru. Repositori berfungsi sebagai lokasi penyimpanan terpusat untuk data – data yang diteliti dan proses analisa maupun hasil dari penelitian ini.



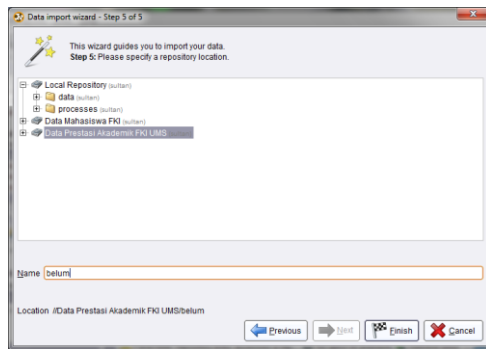
Gambar 3. Pembuatan Repositori

Kemudian masukkan data yang akan dijadikan sebagai *data training*. Dalam penelitian ini penulis langsung mengambil data yang berasal dari *Excel sheet* untuk mengklasifikasikan prestasi akademik dari mahasiswa FKI yang sudah lulus untuk dijadikan *data training*.



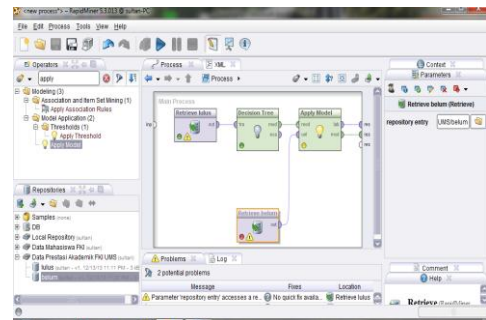
Gambar 4. Import File Excel

Langkah selanjutnya adalah proses pengambilan data uji pada Rapid Miner 5. Sama seperti dalam proses pengambilan *data training*.



Gambar 5. Penyimpanan file data uji
Setelah mempunyai *data training* dalam hal ini data lulusan mahasiswa FKI UMS dan data uji yaitu data yang belum lulus dari mahasiswa FKI UMS, data tersebut kemudian diuji menggunakan klasifikasi *decision tree* dengan cara *drag and drop* kedalam halaman *main process* pada tampilan *design perspektif*.

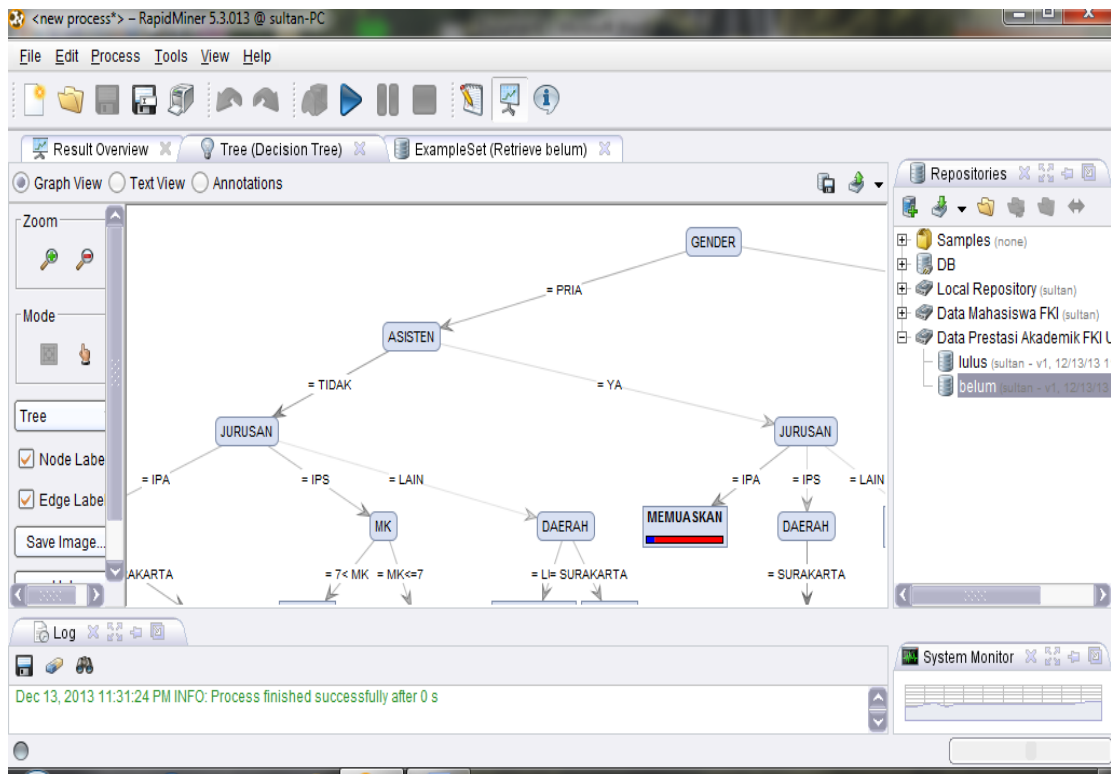
Adapun model – model yang di *drag and drop* ke halaman *main process* adalah *data training*, data uji, model klasifikasi yaitu *decision tree* serta *apply model*.



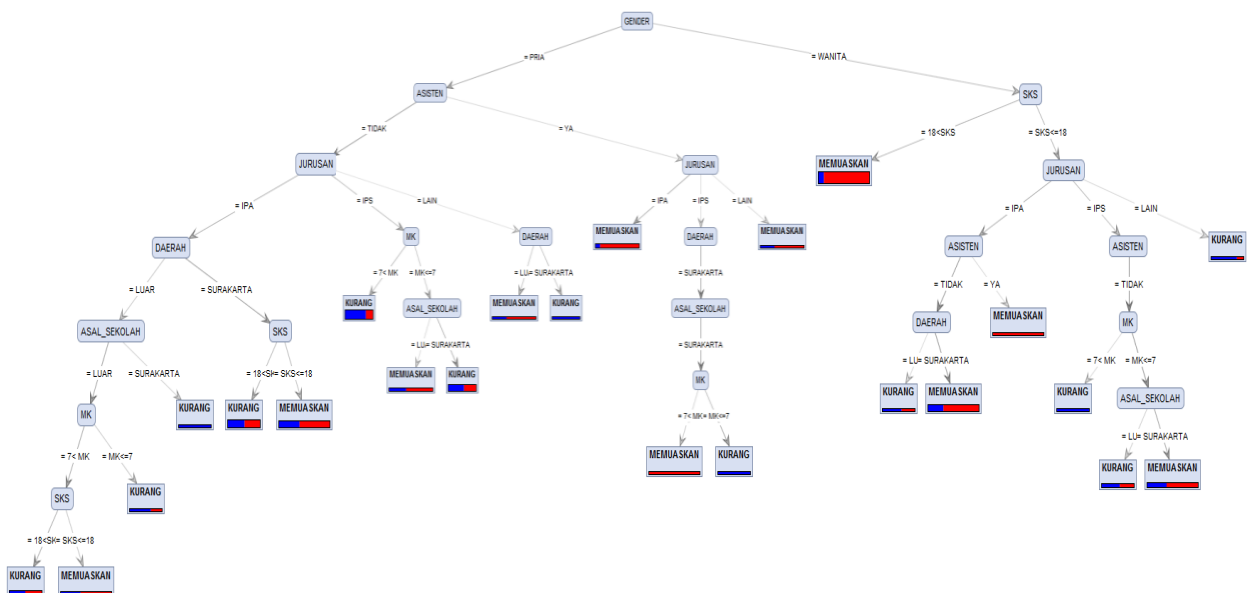
Gambar 6. Penghubungan setiap *port* model

Setelah semua terhubung kemudian klik tombol *run*, sehingga pada *result perspective* akan muncul pohon keputusan (*decision tree*) dari data – data yang dimasukkan.

Hasil klasifikasi dengan aplikasi Rapid Miner menunjukkan bahwa data akademik FKI UMS yang semula kurang bermanfaat dapat diolah berdasarkan atribut - atribut yang dibutuhkan menjadi data yang bermanfaat bagi fakultas, dibuktikan dengan hasil klasifikasi pada penelitian ini. Atribut yang paling berpengaruh adalah gender. Dalam *graph view* terlihat bahwa yang diprediksi memiliki IPK memuaskan berasal dari jurusan IPA serta juga terlihat bahwa yang gender pria dan tidak menjadi asisten, dari wilayah karesidenan Surakarta diprediksi memiliki IPK yang kurang memuaskan.



Gambar 7. Hasil *decision tree* tampilan pohon keputusan.



Gambar 8. Hasil keseluruhan dari *decision tree*.

Dari hasil *decision tree* memperlihatkan bahwa yang berjenis kelamin wanita dan berasal dari luar karesidenan Surakarta juga diprediksi memiliki IPK yang kurang memuaskan, sehingga dengan hasil penelitian bisa memberikan rencana strategis untuk FKI UMS dalam merekrut calon mahasiswa baru bisa menggunakan skala prioritas dari jurusan IPA, banyak merekrut calon mahasiswa pria dari luar karesidenan Surakarta serta membatasi pengambilan calon mahasiswa perempuan dari luar karesidenan Surakarta.

V. PENUTUP

1. Kesimpulan

Berdasarkan *training* dan pengujian dapat diambil beberapa kesimpulan sebagai berikut:

Data mining dengan teknik klasifikasi pada data mahasiswa FKI UMS menggunakan metode *decision tree* menghasilkan informasi bahwa hasil klasifikasi menunjukkan bahwa beberapa atribut mempunyai dominasi yang cukup tinggi terhadap nilai IPK mahasiswa seperti jurusan sekolah di

SMA, yang paling berpengaruh adalah IPA. Informasi yang ditampilkan berupa nilai *support* dan *confidence* dari hubungan antara IPK mahasiswa dengan data induk mahasiswa. Dari pohon *tree* yang terbentuk dapat dilihat bahwa mahasiswa yang diprediksi mempunyai IPK memuaskan adalah :

- a. Mahasiswa berjenis kelamin pria dari karesidenan Surakarta non IPA diprediksi memiliki IPK kurang dan dari luar Surakarta diprediksi memiliki IPK memuaskan.
- b. Mahasiswa jenis kelamin wanita dari karesidenan Surakarta diprediksi memiliki IPK memuaskan, sedang dari luar Surakarta diprediksi memiliki IPK kurang memuaskan baik dari jurusan IPA maupun IPS.
- c. Seluruh mahasiswa dari jurusan IPA mayoritas diprediksi memiliki IPK yang memuaskan.

2. Saran

Berdasarkan kesimpulan yang di dapat maka dapat diberikan saran:

- a. Data mahasiswa baru sebaiknya dikelola dengan baik karena dapat membawa manfaat yang sangat besar untuk FKI UMS.
- b. Perlu adanya penelitian lebih lanjut berdasarkan atribut-atribut lain yang belum dimanfaatkan seperti kegiatan ekstra kurikuler mahasiswa serta rata – rata indeks prestasi per semester.
- c. Pengoptimal peran dosen wali pada saat konseling input KRS. Terutama mahasiswa yang berasal dari jurusan selain IPA serta jenis kelamin wanita dari luar karesidenan Surakarta diperlukan konseling yang lebih intensif dan tepat.
- d. Dalam penerimaan mahasiswa baru, menurut *decision tree* yang terbentuk diprioritaskan untuk memperbanyak calon mahasiswa dari jurusan IPA, jika wanita yang berasal dari karesidenan Surakarta.

DAFTAR PUSTAKA

- Hastuti, Khafiizh (2012). *Analisis Komparasi Algoritma Klasifikasi Data Mining Untuk Prediksi Mahasiswa Non Aktif*. Semarang : Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2012. ISBN 979 - 26 - 0255 - 0.
- Herman, Aldino dan Naam, Julfriadi (2012). *Penerapan Kata Kunci Dokumen Teks Surat Kabar Singgalang Dengan Menggunakan Metoda Algoritma Naive Bayes*. Singgalang.
- Prasetyo, Bowo (2011). *Tool Open Source untuk Data Mining*. From : <http://www.slideshare.net/bowoprasetyo/rapidminer>, diakses tanggal 16 Desember 2013 jam 16.00 WIB.
- Santosa, Budi (2007). *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta : Graha Ilmu.
- Satrio Wahono, Romi (2012). *Proses Data Mining*. From : <http://www.romisatriawahono.net/lecture/dm/romi-dm-02-proses-june2012.pptx>, diakses tanggal 16 Desember 2013 jam 15.30 WIB.
- Umar, Husein (2004). *Metode Penelitian Untuk Skripsi Dan Tesis Bisnis*. Cetakan ke-6. Jakarta: PT Raja Grafindo Persada.
- Wahyu Ningrat, Ranny. Santosa, Budi (2012). *Pemilihan Diet Nutrien bagi Penderita Hipertensi Menggunakan Metode Klasifikasi Decision Tree (Studi Kasus: RSUD Syarifah Ambami Rato Ebu Bangkalan)*. Surabaya :Teknik Industri, Fakultas Teknologi Industri, Institut Teknologi Sepuluh Nopember (ITS). JURNAL TEKNIK ITS Vol. 1, No. 1, ISSN: 2301-9271 A-536.

BIODATA PENULIS

Nama : Setyawan
Tempat / Tanggal Lahir : Klaten, 05 Juli 1976
Jenis Kelamin : Pria
Agama : Islam
Jurusan : Teknik Informatika
Perguruan Tinggi : Universitas Muhammadiyah Surakarta
Alamat : Jl. A. Yani Tromol Pos I Pabelan, Kartasura
Telp./ Fax : (0271)717417, 719483 / (0271) 714448
Alamat Rumah : Plelen RT.04/33 Kadipiro, Banjarsari, Surakarta
No. HP. : +6285329004171
Alamat Email : wanzetya@gmail.com